By ALBERT A. HIRSCH

The BEA Quarterly Model as a Forecasting Instrument

SINCE the early 1960's, there has been an explosive growth in the number of econometric models of the U.S. economy. The models now existing differ considerably as to the underlying theoretical framework, intended use, extent of disaggregation, specific formulation of comparable relationships, and time unit (month, quarter, or year).

The Bureau of Economic Analysis (BEA) quarterly model belongs to the family of "large" models whose primary use is in forecasting. It is also used to analyze the impact of alternative Government policies. It is a substantially revised and enlarged version of a model published in 1966.[1] The model now contains 63 behavioral equations. The model equations and an explanation of them are presented in a BEA staff paper.[2] An interim version of the model, together with an analysis of sample period predictions and cyclical properties, is presented elsewhere.[3]

This article examines and evaluates the predictive and forecasting properties of the model. In the present context, the following distinction is drawn between prediction and forecasting. Prediction pertains to the determination of values of the "endogenous" variables in the model (the variables that are explained by the system of equations), when the values of the "exogenous" variables (those that are determined outside the system) and the "initial conditions" (values of both kinds of variables in the period(s) just before the prediction period) are "given" (in the sense of being known or being hypothetically assumed). Predictions can refer to the past as well as to the future. One may also assume both the initial conditions and the exogenous variable hypothetically, in which case the predictions do not pertain to any historical time period at all. The predictions that are analyzed in this article, however, were made with actual initial conditions and known exogenous variables.

Forecasting refers to the estimation of the probable future values of economic variables. Forecasts differ from predictions in that they require judgmental forecasts of exogenous variables in place of known or hypothetical values. (As is explained later, most econometric forecasters introduce other judgmental elements as well.) They share in common with predictions the feature that the model translates the initial conditions and exogenous variables into a set of endogenous outputs. The end result in the case of forecasts, however, is a set of unconditionally stated expectations of what is most likely to happen, thus contrasting with the conditional nature of predictions. (In practice, we frequently deviate from forecasting in this pure sense by generating more than one set of outputs for a given time period, based on alternative assumptions for some of the inputs.)

The first major section of the article presents an investigation of the errors made when the model is used without judgmental modifications to predict both within and beyond the "sample period," that is, the period from which the data used to fit the equations are taken. The basic summary statistic used to measure error is the "root mean square error," which is defined later. While the main focus is on the quantitative accuracy of predictions, the degree of success in predicting business cycle turning points is also examined. The second major section analyzes the record of errors in actual forecasts made with various versions of the BEA model over the period 1966–71 and compares those results with the post-sample prediction errors and with errors generated by certain other procedures. A final section summarizes the major statistical findings.

The whole inquiry—both its prediction and forecasting aspects—is aimed at the question: How reliable is the model as a forecasting instrument? The article does not provide an unambiguous answer to this question. However, both the quantitative error statistics and the analysis of turning point predictions show a substantial tendency toward deterioration as the prediction or forecast horizon lengthens. Since a large part of the impact of many kinds of Government economic policy actions occurs several quarters after such actions, further improvements in econometric modeling are desirable.

An econometric model is a set of equations comprised of behavioral relationships plus "identities," or definitional relationships. The behavioral relationships are specified (as far as possible) on the basis of economic theory and are estimated by fitting regressions to actual data. A basic assumption is that the relationships are "stochastic." That is, even if all of the important causal determinants are

---

1. Maurice Liebenberg, Albert A. Hirsch, and Joel Popkin, "A Quarterly Econometric Model of the United States: A Progress Report," SURVEY OF CURRENT BUSINESS, May 1966. The model was generally referred to as the "OBE model" prior to the renaming of the Office of Business Economics as the Bureau of Economic Analysis.

2. Albert A. Hirsch, Maurice Liebenberg, and George R. Green, "The BEA Quarterly Model," Bureau of Economic Analysis Staff Paper No. 22, July 1973, available from the National Technical Information Service, Springfield, Virginia 22151. Order by COM 73-11114. Price is $3.00.

3. George R. Green, in association with Maurice Liebenberg and Albert A. Hirsch, "Short- and Long-term Simulations with the OBE Econometric Model" in *Econometric Models of Cyclical Behavior*, edited by Bert G. Hickman, Studies in Income and Wealth, 36, Vol. 1, National Bureau of Economic Research, 1972.

NOTE.—Computational and clerical assistance in the preparation of this article by Fannie M. Hall, Irene M. Mattis, and Judith K. White is gratefully acknowledged.

included as explanatory variables in an equation and the form of the equation is properly specified, there remains a random or unexplained error term (often called "disturbance") which represents the net effect of the myriad other forces that are acting on the dependent variable.

The stochastic nature of the behavioral relationships is a fundamental source of prediction error in an econometric model. The unexplained, and therefore unpredictable, random disturbances are a direct cause of errors in the dependent (endogenous) variables. They also cause prediction error indirectly since in the estimation process they give rise to sampling errors in the coefficients; and erroneous coefficients result in wrong predictions of the effects that changes in one variable have upon another.

A second source of prediction error is errors in specifying behavior relationships. Specification errors can take the form of omission of important causal variables that should be included, inclusion of variables that should be excluded, or incorrect mathematical formulation of the function. A related source of error which manifests itself primarily in post-sample predictions and forecasts is structural change, i.e. the tendency for the "true" parameters of the system to change over time. It can be argued that this phenomenon simply reflects omission of variables, but such omissions can be manifold and hard to pinpoint or quantify.

A third class of sources of prediction error are various special problems of estimation. As in the case of random disturbances, these problems can give rise to prediction error via their tendency to yield incorrect estimates of parameters. These include intercorrelation among explanatory variables, autocorrelation of disturbances, simultaneity of equations, inclusion of lagged dependent variables, errors in the measurement of data, and again, errors in specification. The deficient characteristics of the estimated coefficients that arise when appropriate corrective econometric techniques are not or cannot be applied are mainly "bias" (the tendency to under- or overestimate the parameters), and "inefficiency" (the tendency for parameter estimates based on different samples of data to be widely dispersed).

Forecasts are subject to all the sources of prediction error, but they are also subject to additional sources of error. First, there are errors in projecting exogenous variables. Second, there are errors in the judgmental adjustments made to the model equations. Third, forecasts are normally made using preliminary (and sometimes incomplete) data for initial conditions; this is a source of error that would be absent if complete revised data were used.

Fortunately, the law of large numbers in statistics leads us to expect that the various sources of error tend to be offsetting in their impact on the prediction of particular variables. As the evidence presented in this article shows, there is an analogous offsetting tendency with respect to errors in the components of key aggregates.

Prediction Errors

There is no direct way to test the forecasting ability of a model prior to actual forecasting use. However, since prediction errors are likely to be major contributors to forecasting errors, an obvious indirect test of a model's likely forecasting performance is to see how well the model predicts endogenous variables in periods for which actual values are known, using known values of exogenous variables, the latest revised data for the initial conditions, and perhaps crude estimates of the direct impacts of such exogenous factors as major strikes or strike threats. Such tests can be made both within and beyond the sample period.

Tests beyond the sample period are particularly crucial since these indicate the stability of the model relationships; good performance within the sample period may reflect, in part, *ad hoc* selection of relationships on the criterion of fit. Unfortunately, the "degrees of freedom" provided by usable time series observations are relatively scarce. Accordingly, most of the available observations must be used for fitting the model's equations, leaving only a small number of available periods for post-sample testing. For this reason, as well as the need to determine how well the post-sample performance holds up relative to that of the sample period, the error statistics for the latter provide important additional information.

One- to six-quarter predictions

The quantitative prediction errors were obtained by running dynamic simulations for overlapping six-quarter spans covering both the bulk of the sample period, which is 1953–II—1968–IV, and a post-sample period (1969–I—1971–II) and comparing predicted with actual values. In dynamic predictions, computed rather than actual values of lagged endogenous variables are used as inputs for subsequent periods. The first sample-period prediction sequence begins in 1953–IV and ends in 1955–I; the second begins in 1954–I and ends in 1955–II, and so on, through the sequence ending in 1968–IV. A six-quarter span was chosen because that is a usual forecasting horizon. Error statistics (to be described shortly) were calculated for all predictions one quarter ahead, all predictions two quarters ahead, . . ., and all predictions six quarters ahead. In order to calculate all the summary error statistics with the same number of observations for all horizons and to have all of them cover the same time period, errors for periods prior to 1955–I were not used. Thus, the sample-period error statistics are based on 56 sets of predictions covering the 14 years from 1955–I through 1968–IV. For the post-sample period, there are only 10 sets of predictions covering the 2½ years from 1969–I through 1971–II.[4]

In making the predictions, account was taken of serial correlation in the equation residuals, that is, the tendency for the residuals of successive time periods to be systematically related; this is, after all, useful information that should not be discarded in making pre-

4. The data reported in this study were compiled prior to the July 1972 revision of the national income and product accounts. Hence the "actual" values of national income variables against which errors are measured are based on the accounts as of July 1971.

dictions. Specifically, in equations in which serial correlation is significantly present, additive adjustments were made to the constant term based on the last two observed residuals prior to the prediction period and the estimated autocorrelation coefficient. The adjustments are such that they decay from a weighted average of the two residuals toward zero over the prediction period. In algebraic terms:

$$\mathrm{Adj}_{t+i} = \tfrac{1}{2}\, b^{i}(e_t + b e_{t-1})$$

where e_t is the observed residual in the initial quarter (first quarter prior to prediction period), e_{t-1} is the residual in the previous quarter, b is the estimated autocorrelation coefficient, and i is the number of quarters from the initial quarter being predicted.[5]

Size of prediction errors

The basic summary error statistic in this study is the root mean square error (RMSE), which is given by the formula

$$\frac{1}{N}\sum_{i=1}^{N}(P_i - A_i)^2$$

where P_i represents the predicted value for the i-th observation, A_i the corresponding actual, and N the number of observations. Alternative measures, such as the average absolute error,[6] could be used. However, the RMSE has useful analytical characteristics; in particular, its square (the mean square error) can be decomposed into contributing elements, as shown later, whereas the average absolute error cannot. It should be kept in mind, when evaluating RMSE statistics, that the RMSE gives more weight to extreme errors than does the average absolute error and thus tends to be larger.

Sample period. Table 1 shows RMSE's for the period 1955–I—1968–IV for predictions of major variables with horizons of from one to six quarters. Only RMSE's for endogenous components of GNP are shown since exogenous variables are assigned their actual values and hence show no error; for this reason, exports, military imports, and government purchases of goods and services are not listed. Second, it should be noted that the model determines components of real GNP and corresponding price deflators; thus the RMSE's for current-dollar magnitudes—real magnitudes times prices—represent composites of errors in the basic variables.

Two generalizations can be made about the RMSE's. First, the errors generated grow in size as the prediction horizon lengthens. This phenomenon reflects accumulation of errors through lagged variables, which, after the first predicted quarter, also contain prediction errors. The tendency toward increasing error is greatly subdued in the prediction of quarterly changes, as can be seen in the second line of table 1 where RMSE's for change in current-dollar GNP are shown. The reason for this is that in any prediction sequence the accumulation of errors through lagged variables tends to be in one direction; to this extent, accumulation is registered in the levels, but not in the changes. The second generalization is that RMSE's for aggregates, such as GNP or personal consumption expenditures, are less than the sum of component RMSE's. This reflects the tendency of errors to be offsetting.

The largest errors among components of real GNP are in nondurables consumption, nonresidential fixed investment, and change in business inventories. (The size of the RMSE's in nondurables consumption is not surprising since this is the largest single component of real final demand in the model.) Relative sizes of errors in the real final demand components are roughly reflected in those of the corresponding current-dollar magnitudes.

RMSE's for personal income are larger than those for corporate profits, and increase more rapidly as the prediction horizon lengthens. However, profits are much smaller than personal income; thus, in percentage terms (not shown), profit predictions are subject to considerably larger errors.

Compared with errors in current-dollar GNP and real GNP, errors in the implicit price deflator for private GNP are surprisingly small. This can be seen from comparisons of root mean square percentage errors for the three magnitudes, as shown in the last three lines of table 1. A percentage error is computed as

$$\frac{P-A}{A}\times 100,$$

and the root mean square percentage error is calculated analogously to the RMSE. Because errors in the price level predictions are relatively small, errors in real GNP carry through directly into errors in current-dollar GNP. In the first prediction quarter, the root mean square percent error in real GNP is more than three times as great as that in the private deflator. By the fifth and sixth prediction quarters this ratio is nearly four. However, the relatively small errors in the aggregate price index reflect larger but offsetting errors among component price deflators.

Errors in the unemployment rate reflect, in part, those in real output. This is seen in the fact that errors in employment, which is directly related to output, are larger (the more so as the prediction horizon lengthens) than those in labor force. Errors in both short- and long-term interest rates remain quite low over the whole six-quarter prediction horizon.

Post-sample period. Error statistics for predictions beyond the sample period are derived from only 10 sets of overlapping predictions covering 2½ years, as against 56 full sets covering 14 years in the sample period. Predictions over six-quarter spans in the post-sample period (1969–I through 1971–II) were obtained in the same way as the sample-period predictions with two modifications: (1) The constant adjustments were made to decay over the prediction horizon from the average of the last

---

5. The adjustments were made to the "normalized" equations. Thus, if the dependent variable of a behavioral equation that is estimated in constructing the model is not a simple endogenous variable, but, for example, a ratio such as CSNH/N (nonhousing services consumption per capita), then the equation is first transformed so that only a single endogenous variable appears on the left. In the example here, both sides of the equation are first multiplied by N (population) before the adjustment formula is applied.

It should be noted that use of the above formula for the sample period predictions is quantitatively not very important, since many of the estimated equations already contain a correction for serial correlation in the residuals via the "Cochrane-Orcutt" transformation. Thus, with a few exceptions, the b coefficients are relatively small.

6. Average absolute error =

$$\frac{1}{N}\sum_{i=1}^{N}|P_i - A_i|$$

i.e., in computing the sum of the errors, the signs of individual errors are disregarded.

two periods' residuals prior to the prediction period toward an average of the last eight quarters' residuals rather than toward zero;[7] (2) special adjustments were incorporated to handle the direct effects of the General Motors strike in late 1970. The first modification allows for the fact that beyond the sample period, the average prediction error of the equation may differ from zero because of specification errors or gradual structural change. The eight-quarter average error is intended to represent an updated long-run average of expected errors. The second modification takes into consideration that the impact of the GM strike was considerably greater than that of previous strikes, whose mean effects are represented by the coefficients of strike dummy variables.

Prediction errors for the post-sample period, shown in table 2, are generally larger than for the sample period. This result is to be expected. Most of the variables in the model exhibit substantial growth trends. It is thus natural that the prediction errors should be larger in the post-sample period when the values of the variables are large as compared with those of the sample period. Furthermore, it can be shown that even if the random disturbances do not increase in size with that of the endogenous variables, expected prediction errors grow with the increasing gap between current and sample mean values of the explanatory variables. The tendency toward increasing error is aggravated by errors in specification and by structural changes. Another likely reason for the larger errors is that this particular period was an inherently difficult one to predict.

For current-dollar GNP, the ratio of RMSE's in the 1969–71 period to RMSE's in the 1955–68 period averages to about 2.3 over the whole prediction horizon (see table 3). For real GNP this ratio is about 1.7. For the private GNP deflator, the ratio is 1.5 for one-quarter predictions and rises to 2.6 for six-quarter predictions. The "amplification" of the root mean square percentage errors is much smaller, though for each variable and all prediction horizons it is still greater than 1.0.

Among GNP components, errors in personal consumption expenditures are much larger relative to errors in total GNP in the post-sample than in the sample period. Most of this difference is accounted for by substantially larger

7. The formula used is
$$Adj_{t+i}=\tfrac{1}{2}b^{i}[(e_{t}-\bar{e}_{t-8})+b(e_{t-1}-\bar{e}_{t-8})]+\bar{e}_{t-8},$$
where $\bar{e}_{t-8}=\tfrac{1}{8}\sum_{j=1}^{8}e_{t-j+1}$.

Table 1.—Root Mean Square Errors of Selected Variables: Sample Period Predictions (1955-I—1968-IV)

Table 2.—Root Mean Square Errors of Selected Variables: Post-sample Period Predictions (1969-I—1971-II)

Table 3.—Ratios of Post-Sample Period to Sample Period Root Mean Square Errors

	Table 1: Prediction horizon (quarters)						Table 2: Prediction horizon (quarters)						Table 3: Prediction horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q	1Q	2Q	3Q	4Q	5Q	6Q	1Q	2Q	3Q	4Q	5Q	6Q
Billions of dollars:																		
Gross national product	2.94	5.16	7.31	9.36	11.61	12.65	6.01	11.01	18.42	23.63	28.08	30.56	2.0	2.1	2.5	2.5	2.4	2.2
Change in GNP	2.94	3.89	3.93	4.06	4.10	4.25	6.01	7.43	7.46	6.79	6.68	6.48	2.0	1.9	1.9	1.7	1.6	1.5
Personal consumption expenditures	1.91	2.77	3.66	4.63	5.90	7.10	3.98	6.97	10.89	13.77	16.69	18.82	2.1	2.5	3.0	3.0	2.8	2.7
Fixed nonresidential investment	.95	1.69	2.36	3.04	3.79	4.50	2.05	3.41	4.95	6.66	8.84	10.71	2.2	2.0	2.1	2.2	2.3	2.4
Residential structures	.54	1.06	1.43	1.66	1.73	1.72	1.26	2.88	3.56	3.64	3.48	3.31	2.3	2.7	2.5	2.2	2.0	1.9
Change in business inventories	2.03	2.59	2.82	3.12	3.46	3.74	3.01	3.66	4.54	5.64	5.31	4.83	1.5	1.4	1.6	1.8	1.5	1.3
Imports of goods and services	.51	.72	.98	1.12	1.28	1.40	3.56	3.84	4.03	4.99	5.21	5.44	7.0	5.3	4.4	4.5	4.2	3.8
Personal income	1.97	3.14	4.22	5.36	6.72	8.16	3.28	7.70	13.33	18.00	22.63	25.75	1.7	2.5	3.3	3.4	3.4	3.2
Corporate profits and inventory valuation adjustment	1.95	2.84	3.78	4.60	5.22	5.71	2.18	3.20	4.86	5.82	5.60	5.83	1.1	1.1	1.3	1.2	1.1	1.0
Billions of 1958 dollars:																		
Gross national product	2.80	4.93	6.75	8.46	10.25	11.53	4.56	8.11	13.28	16.08	17.78	17.16	1.6	1.7	2.0	1.9	1.7	1.5
Personal consumption expenditures	1.90	2.88	3.61	4.30	5.16	5.93	3.21	6.60	7.98	9.41	10.31	10.11	1.7	2.1	2.3	2.2	2.0	1.7
Automobiles and parts	.77	1.00	1.20	1.42	1.62	1.74	1.79	2.46	3.20	3.88	4.19	4.29	2.3	2.5	2.7	2.7	2.6	2.5
Nonauto durables	.48	.59	.67	.78	.89	1.01	1.08	1.57	2.04	2.31	2.28	1.90	2.2	2.7	3.0	3.0	2.5	1.9
Nondurables	1.32	1.56	1.96	2.10	2.86	2.80	1.78	2.26	2.83	3.10	3.38	3.16	1.3	1.5	1.5	1.5	1.4	1.2
Nonhousing services	.39	.45	.54	.58	.56	.63	1.26	1.38	2.32	2.68	2.88	2.90	3.3	3.8	4.3	4.7	5.0	4.6
Housing services	.26	.45	.60	.78	1.00	1.23	.17	.20	.30	.28	.54	.67	.7	.5	.5	.5	.5	.5
Fixed nonresidential investment	.87	1.52	2.11	2.70	3.38	4.02	1.50	2.66	3.77	4.77	6.36	7.66	1.7	1.8	1.8	1.8	1.9	1.9
Residential structures	.52	.99	1.31	1.49	1.52	1.45	.91	2.09	2.65	2.67	2.48	2.30	1.8	2.1	2.0	1.7	1.6	1.6
Change in business inventories	1.90	2.48	2.71	3.00	3.31	3.60	2.52	3.06	3.79	4.85	4.48	3.91	1.3	1.2	1.4	1.6	1.4	1.1
Merchandise imports	.46	.60	.88	1.08	1.19	1.34	2.97	3.10	3.07	3.72	3.76	3.77	6.5	4.6	3.6	3.6	3.3	2.8
Services imports, nondefense	.13	.14	.15	.16	.17	.18	.76	.93	1.06	1.16	1.21	1.23	5.8	6.6	7.1	7.3	7.1	6.8
Miscellaneous variables:																		
Implicit price deflator, private GNP (1958=100)	.18	.30	.40	.48	.52	.58	.27	.44	.61	.70	.95	1.51	1.5	1.5	1.5	1.5	1.8	2.6
Wages per private employee (dollars per year)	30.4	43.6	57.2	68.7	77.4	89.0	50.0	74.1	110.2	143.9	175.6	200.2	1.6	1.7	1.9	2.1	2.3	2.2
Index of private output per manhour (1958=100)	.72	.82	1.04	1.20	1.26	1.52	.93	1.36	1.44	1.71	1.89	1.87	1.3	1.7	1.6	1.4	1.4	1.2
Civilian labor force (millions)	.24	.28	.30	.32	.34	.35	.42	.50	.74	.80	1.08	1.14	1.8	2.1	2.6	2.8	3.0	3.3
Employed	.30	.40	.43	.49	.56	.66	.42	.78	1.15	1.52	1.88	2.02	1.4	2.0	2.7	3.1	3.2	3.1
Unemployment rate (percent)	.32	.48	.58	.65	.73	.70	.22	.80	.67	.93	1.15	1.26	.7	.8	1.2	1.4	1.6	1.6
Average yield on 4-6 months commercial paper (percent)	.18	.23	.24	.24	.25	.27	.70	.86	.84	.84	.95	1.11	3.9	3.7	3.5	3.5	3.8	4.1
Average yield, corporate bonds (percent)	.09	.13	.16	.18	.20	.20	.38	.49	.49	.47	.46	.48	4.2	3.8	3.3	2.6	2.3	2.4
Root Mean Square Percentage Errors:																		
Gross national product	.51	.88	1.25	1.59	1.98	2.34	.64	1.19	2.00	2.53	2.99	3.28	1.3	1.4	1.6	1.6	1.5	1.4
GNP in 1958 dollars	.54	.92	1.28	1.59	1.90	2.19	.68	1.12	1.83	2.22	2.45	2.87	1.2	1.2	1.4	1.4	1.3	1.1
Implicit deflator, private GNP	.17	.28	.36	.44	.49	.57	.21	.34	.47	.56	.78	1.20	1.2	1.2	1.3	1.3	1.6	2.1

errors in purchases of autos and parts and of nonhousing services. Errors in imports are also much larger. By contrast, errors in housing services are smaller and errors in inventory investment are only moderately larger.

Errors in personal income are much larger. This results from positive bias (discussed below) both in the average private wage rate and in private employment, which result in large errors in employee compensation. The errors in corporate profits, however, are only slightly larger.

The unemployment rate is very well predicted on the average in one- and two-quarter predictions, but the error mounts more rapidly in subsequent quarters than in the comparable quarters within the sample period. Errors in the short- and long-term interest rates are uniformly larger.

Bias component of errors

To what extent are prediction errors the result of systematic factors rather than purely random? There are several kinds of systematic error, the most important of which is bias, that is, a persistent tendency to underpredict or overpredict.[8] The degree of bias in sample period and post-sample period predictions is examined here.

The quantitative importance of bias can be analyzed by decomposing the mean square error (MSE) into the bias component $(\bar{P}-\bar{A})^2$—that is, the square of the average prediction error—and the variance of the error around the average (S^2_{P-A}):

$$MSE = RMSE^2 = (\bar{P}-\bar{A})^2 + S^2_{P-A}$$

Table 4 shows, for sample-period predictions of selected variables, the average prediction error $(\bar{P}-\bar{A})$, the standard deviation of the mean prediction error $(S_{\bar{P}-\bar{A}})$, and the "bias proportion" of the MSE, that is $(\bar{P}-\bar{A})^2$/MSE. The average error is a direct measure of bias which preserves its sign (direction) and the bias proportion indicates the importance of bias in the total error. Table 5 shows the same statistics for post-sample predictions.

During the sample period, average errors for current-dollar GNP and real GNP are small for the whole prediction horizon and not statistically significant.[9] There is also an absence of significant bias in most GNP components, personal income, corporate profits, the implicit deflator for private GNP, and the unemployment rate. There is, however, a significant negative bias in imports for all six quarters. There is also a noticeable positive bias in business inventory investment, but it is not significant at the 5 percent level. Except for imports, the bias proportion is well under 10 percent.

The post-sample period errors present a marked contrast to those of the sample period in respect to bias. (However, the caution given in footnote 9 about interpreting the significance test for bias applies even more strongly to the post-sample than the sample period because of the much smaller number of observations in the former.)

Average errors in both current-dollar GNP and real GNP are positive and, after the first quarter, significantly so at the 5 percent level. Moreover, bias accounts for a sizable proportion of the mean square error. Average errors in the private GNP deflator are also positive, but are significant only in the fifth and sixth quarters.

All major endogenous GNP components begin to show significant positive bias at some point within the six-quarter prediction horizon. For residential construction and business inventory investment, the bias is significant from the start; for other components, it is significant only after the second or third quarter. The positive bias in imports tends to dampen the positive bias in GNP.

On the income side, there is significant positive bias in personal income. There is negative bias in the unemployment rate and significant positive bias in the private GNP deflator after the fourth quarter.

Turning point errors

Thus far we have been concerned with the quantitative aspect of predictive performance, that is, with the size of errors. Of perhaps equal importance is the ability of a model to detect well in advance changes in the direction of economic activity. The degree of reliance that can be placed on models to anticipate business cycle turning points depends on the extent to which they incorporate the cyclical dynamics of the real world.

The view of the business cycle that is consistent with the structure of most econometric models is that the basic (nonstochastic) behavior relationships in the economy do not result in sustained cycles, but that cycles are induced by interaction, via dynamic lag patterns, between that system of relationships, on the one hand, and random shocks to those relationships and to the smooth paths of the exogenous variables, on the other. The theoretical foundation for this view was developed by Slutsky.[10] Its relevance was later tested first on an earlier annual U.S. model [11] and more recently on two quarterly models including a version of the BEA model.[12] In the first study a clear similarity was found between observed historical cycles and cycles simulated by a model "shocked" with random disturbances. In the more recent study, the similarity was found to be somewhat more tenuous.

Nonstochastic simulations with the BEA model and with two other quarterly models, made continuously over each model's full sample period and beyond (i.e., simulations without

8. Two other kinds of systematic error may be noted. First, even if errors are on the average unbiased, it is possible that for low values of a variable actuals are underpredicted while for high values they are overpredicted, or vice versa. Such predictions are said to be inefficient. Another systematic factor frequently found in judgmental forecasts, is underprediction of changes, whether positive or negative. Under estimation of change is not necessarily inconsistent with unbiasedness and efficiency.

9. The 5-percent level of significance is used, based on the t-test, $t=(\bar{P}-\bar{A})/S_{\bar{P}-\bar{A}}$. A t-ratio of approximately 2 or more indicates that the mean error is significantly different from zero, that is, that bias is significant. It should be noted, however, that in the present context the significance test is deficient since the observations are not truly independent because they derive from overlapping predictions.

10. Eugen Slutsky, "The Summation of Random Causes as the Source of Cyclical Processes," *Econometrica*, April 1937.

11. Frank and Irma Adelman, "The Dynamic Properties of the Klein-Goldberger Model," *Econometrica*, October 1959.

12. Victor Zarnowitz, Charlotte Boschan, and Geoffrey H. Moore, "Business Cycle Analysis of Econometric Model Simulations" in *Econometric Models of Cyclical Behavior*, edited by Bert G. Hickman, Studies in Income and Wealth, 36, Vol. 1, National Bureau of Economic Research, 1972.

Table 4.—Bias in Sample Period Predictions of Major Items

[Average errors and standard deviations are in $billions, except as otherwise indicated]

	Prediction horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Gross national product:						
$(P-\bar{X})$[a]	0.53	0.88	0.74	0.49	0.19	−1.01
$S_{P-\bar{X}}$[b]	.50	.88	1.05	1.24	1.42	1.67
$[(P-\bar{X})^2/MSE]$[c]	.532	.499	.010	.093	.000	.096
GNP in 1958 dollars:						
$P-\bar{X}$	.32	.55	.30	−.10	−.76	−1.40
$S_{P-\bar{X}}$	.45	.72	.90	1.07	1.22	1.36
$[(P-\bar{X})^2/MSE]$	.013	.012	.002	.001	.005	.014
Implicit deflator, private GNP (1958=100):						
$P-\bar{X}$	.03	.04	.05	.07	.06	.01
$S_{P-\bar{X}}$	.02	.04	.19	.07	.08	.10
$[(P-\bar{X})^2/MSE]$	.028	.018	.015	.091	.013	.0003
Personal consumption expenditures:						
$P-\bar{X}$	.12	.22	.34	.37	.26	.01
$S_{P-\bar{X}}$	.25	.36	.49	.61	.78	.95
$[(P-\bar{X})^2/MSE]$	.004	.018	.009	.007	.002	.000
Fixed nonresidential investment:						
$P-\bar{X}$	.06	.17	.32	.48	.45	.39
$S_{P-\bar{X}}$	.13	.23	.31	.40	.60	.60
$[(P-\bar{X})^2/MSE]$	.037	.041	.018	.020	.014	.008
Residential structures:						
$P-\bar{X}$	−.05	−.12	−.16	−.20	−.22	−.225
$S_{P-\bar{X}}$	.07	.14	.19	.22	.23	.23
$[(P-\bar{X})^2/MSE]$	.010	.012	.013	.014	.016	.017
Change in business inventories:						
$P-\bar{X}$	.41	.64	.71	.71	.67	.64
$S_{P-\bar{X}}$	.21	.34	.37	.41	.45	.49
$[(P-\bar{X})^2/MSE]$	.041	.061	.064	.052	.128	.029
Imports of goods and services:						
$P-\bar{X}$	−.10	−.20	−.34	−.43	−.49	−.55
$S_{P-\bar{X}}$	.05	.09	.11	.14	.16	.18
$[(P-\bar{X})^2/MSE]$	.081	.131	.130	.141	.143	.144
Personal income:						
$P-\bar{X}$	.36	.65	.79	.69	.39	.01
$S_{P-\bar{X}}$	.31	.49	.62	.78	.86	.99
$[(P-\bar{X})^2/MSE]$	.053	.046	.034	.017	.004	.000
Corporate profits and inventory valuation adjustment:						
$P-\bar{X}$	.23	.30	.28	.25	.10	−.14
$S_{P-\bar{X}}$	.31	.45	.56	.66	.76	.82
$[(P-\bar{X})^2/MSE]$	.024	.018	.008	.003	.000	.000
Unemployment rate (percent):						
$P-\bar{X}$	−.06	−.08	−.08	−.06	−.04	+.00
$S_{P-\bar{X}}$	.04	.06	.07	.07	.06	.06
$[(P-\bar{X})^2/MSE]$	.035	.023	.024	.015	.004	.000

a Average prediction error.
b Standard deviation of average prediction error.
c Bias proportion (square of average error as a proportion of mean square error).

Table 5.—Bias in Post-Sample Period Predictions of Major Items

[Average errors and standard deviations are in $billions, except as otherwise indicated]

	Prediction horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Gross national product:						
$(P-\bar{X})$[a]	1.38	5.10	9.88	12.94	16.81	18.66
$S_{P-\bar{X}}$[b]	.78	1.81	2.12	2.65	3.06	3.24
$[(P-\bar{X})^2/MSE]$[c]	.053	.215	.259	.295	.337	.378
GNP in 1958 dollars:						
$P-\bar{X}$	.48	2.84	5.06	8.31	9.73	9.15
$S_{P-\bar{X}}$	.61	1.09	1.59	1.84	1.99	1.94
$[(P-\bar{X})^2/MSE]$	.011	.141	.204	.308	.332	.250
Implicit deflator, private GNP (1958=100):						
$P-\bar{X}$	.01	.04	.12	.27	.43	1.01
$S_{P-\bar{X}}$	.03	.06	.08	.09	.11	.16
$[(P-\bar{X})^2/MSE]$	.000	.010	.035	.059	.206	.447
Personal consumption expenditures:						
$P-\bar{X}$	−.17	.59	2.30	4.12	6.50	5.90
$S_{P-\bar{X}}$	.53	.92	1.42	1.76	2.05	2.22
$[(P-\bar{X})^2/MSE]$	.002	.047	.045	.090	.152	.234
Fixed nonresidential investment:						
$P-\bar{X}$	.06	.59	1.98	3.30	6.11	6.81
$S_{P-\bar{X}}$	.27	.45	.61	.78	.97	1.11
$[(P-\bar{X})^2/MSE]$	.001	.030	.100	.259	.334	.404
Residential structures:						
$P-\bar{X}$	1.07	2.63	3.41	3.57	3.35	3.14
$S_{P-\bar{X}}$	.09	.15	.13	.10	.13	.14
$[(P-\bar{X})^2/MSE]$	.721	.852	.918	.962	.927	.900
Change in business inventories:						
$P-\bar{X}$	.78	2.08	3.14	3.78	4.00	3.40
$S_{P-\bar{X}}$	.30	.41	.44	.54	.47	.41
$[(P-\bar{X})^2/MSE]$	.058	.300	.478	.461	.667	.551
Imports of goods and services:						
$P-\bar{X}$	.35	.67	1.41	1.87	2.60	3.54
$S_{P-\bar{X}}$	.48	.51	.51	.62	.61	.55
$[(P-\bar{X})^2/MSE]$	.005	.030	.119	.140	.249	.429
Personal income:						
$P-\bar{X}$	.73	3.46	6.95	10.23	13.60	16.39
$S_{P-\bar{X}}$	.42	.93	1.52	1.99	2.42	2.66
$[(P-\bar{X})^2/MSE]$	.030	.197	.273	.323	.361	.404
Corporate profits and inventory valuation adjustment:						
$P-\bar{X}$	.38	1.03	1.30	1.18	1.32	1.06
$S_{P-\bar{X}}$	.28	.41	.63	.70	.73	.73
$[(P-\bar{X})^2/MSE]$	.028	.104	.072	.040	.056	.037
Unemployment rate (percent):						
$P-\bar{X}$	−.01	−.28	−.60	−.75	−.97	−1.08
$S_{P-\bar{X}}$	.03	.05	.06	.10	.12	.12
$[(P-\bar{X})^2/MSE]$	.002	.349	.557	.651	.711	.734

a Average prediction error.
b Standard deviation of average prediction error.
c Bias proportion (square of average error as a proportion of mean square error).

random shocks), show that the models tend to replicate well the first actual business cycle in the period being simulated but to follow only weakly the contours of subsequent cycles, or to miss turning points altogether.[13] This happens in part because the shocks that are reflected in the lagged endogenous variables that define the initial conditions damp out over time, and so, accordingly, does the cyclical behavior of the model since it receives no further shocks other than erratic changes in the exogenous variables.

The above theoretical and empirical evidence leads us to expect that for short prediction horizons models may do reasonably well in predicting turning points. The panels in chart 10 show actual paths of real GNP in the vicinity of specific cyclical turning points in real GNP, and predicted paths using the BEA model. Three six-quarter simulations were run in the vicinity of each turning point. A turning point (downturn or upturn) is defined to be the quarter following a peak or trough in real GNP. The simulations were initiated from one, two, and three quarters before the turning point. The turning points include four upturns beginning with that of 1954 and three downturns beginning with that of 1957. All the recessions and recoveries, except the most recent one, occurred within the sample period.

A summary tabulation showing the degree of the model's success in identifying turning points and the extent of mistimings is given in table 6. As one might expect, the proportion of mistimed turning point predictions increases with the interval between the initial quarter and the turning point. For all predictions, approximately two-thirds of the turning points are correctly predicted. Among the eight cases which do not show the correct timing, predicted turning points are off by more than one quarter in only two cases.

The upturn in 1954-III is well replicated by the model. Each simulation correctly predicts the upturn quarter and follows the actual path of real GNP quite closely. Predictions of the downturn in 1956-IV and the upturn in 1958-II are not nearly as good. While the changes in direction of real GNP are recognized in all six simulations, those beginning more than one quarter before the downturn and that beginning three quarters before the upturn predict the respective turning points one quarter early. More important, the depth of the 1957-58 recession is badly underestimated.

The 1960-II downturn is predicted by all three simulations begun prior to it, despite the relative mildness of that recession. However, the model depicts a shorter recession than actually occurred. In all simulations except the one beginning one quarter before the upturn, an early upturn is predicted.

The moderate downturn in 1969-IV is correctly predicted in each of the downturn simulations. However, the upturn in 1970-II is predicted with a lag by each of the three simulations related to it. The simulation beginning from 1970-I does show a slight gain for 1970-II, but that is followed by two quarters of further decline so that it cannot be regarded as a genuine upturn prediction. Nevertheless, the simulations do follow broadly the contour of the actual economy. (It should be noted that the sharp temporary dip in 1970-IV is associated with the General Motors strike, and is unrelated to the recession; no allowance was made for strike effects in these simulations.)[14]

Forecasting Errors: 1966-71

Forecasts using the BEA quarterly model, beginning with the version published in 1966, have been made regularly with horizons of four or more quarters. From these forecasts summary error statistics have been compiled for the period 1966-I through 1971-II.[15] The historical record does not represent a set of values generated by a constant forecasting mechanism. Rather, it represents the experience of a team using an evolving econometric model. The model structure has been changed continually in both major and minor ways and was periodically reestimated, and each forecast was made with the latest version. There is also a lack of exact consistency among forecasts in the degree of exogeneity and in the choice between expectations-based and endogenous variants of fixed nonresidential investment equations.

Before turning to an examination and evaluation of forecasting performance, we first explain the role of judgmental elements (other than the forecasting of exogenous variables) in model forecasting. Also, two methodological problems involved in compiling forecast errors and their solution are discussed.

Judgmental elements in forecasting

It is possible to generate model forecasts mechanically, just as was done in the case of the predictions, simply by "plugging in" the necessary exogenous variables over the forecasting horizon, perhaps adding formula-based constant adjustments, and solving the model

Table 6.—Prediction of Turning Points in Real GNP

Date of turning point*	Number of quarters from initial quarter to turning point: 1	2	3	All simulations
	Successes (O) or failures (X)			
1954-III (U)	O	O	O	
1957-IV (D)	O	X	X	
1958-II (U)	O	O	X	
1960-II (D)	O	O	O	
1961-II (U)	O	X	X	
1969-IV (D)	O	O	O	
1970-II (U)	X	X	X	
Relative number of successes	6/7	4/7	3/7	13/21
Turning point missed by 1 quarter	0	2	4	6
Turning point missed by 2 quarters	0	1	0	1
Turning point missed by 3 or more quarters	1	0	0	1

*U=Upturn, D=Downturn.

13. *Ibid.*, Section 3.

14. Zarnowitz, Boschan, and Moore (*op. cit.*), examined not only the detection of turning points, but also the degree to which simulations in the vicinity of cyclical turning points replicate, for such business cycle indicators identified by the National Bureau of Economic Research as occur in the models, the lead-lag relationships typical of the actual behavior of those indicators. They conclude that models have a bias toward leads. For the BEA model, they find that "most of the simulated leading and coinciding series lead [actual turns], while lagging series show a tendency to coincide" (*Ibid.*, p. 341). One must emphasize the tentativeness of this finding, however, since the period covered by the model runs is short relative to the time span that forms the basis of the NBER's classification of indicators.

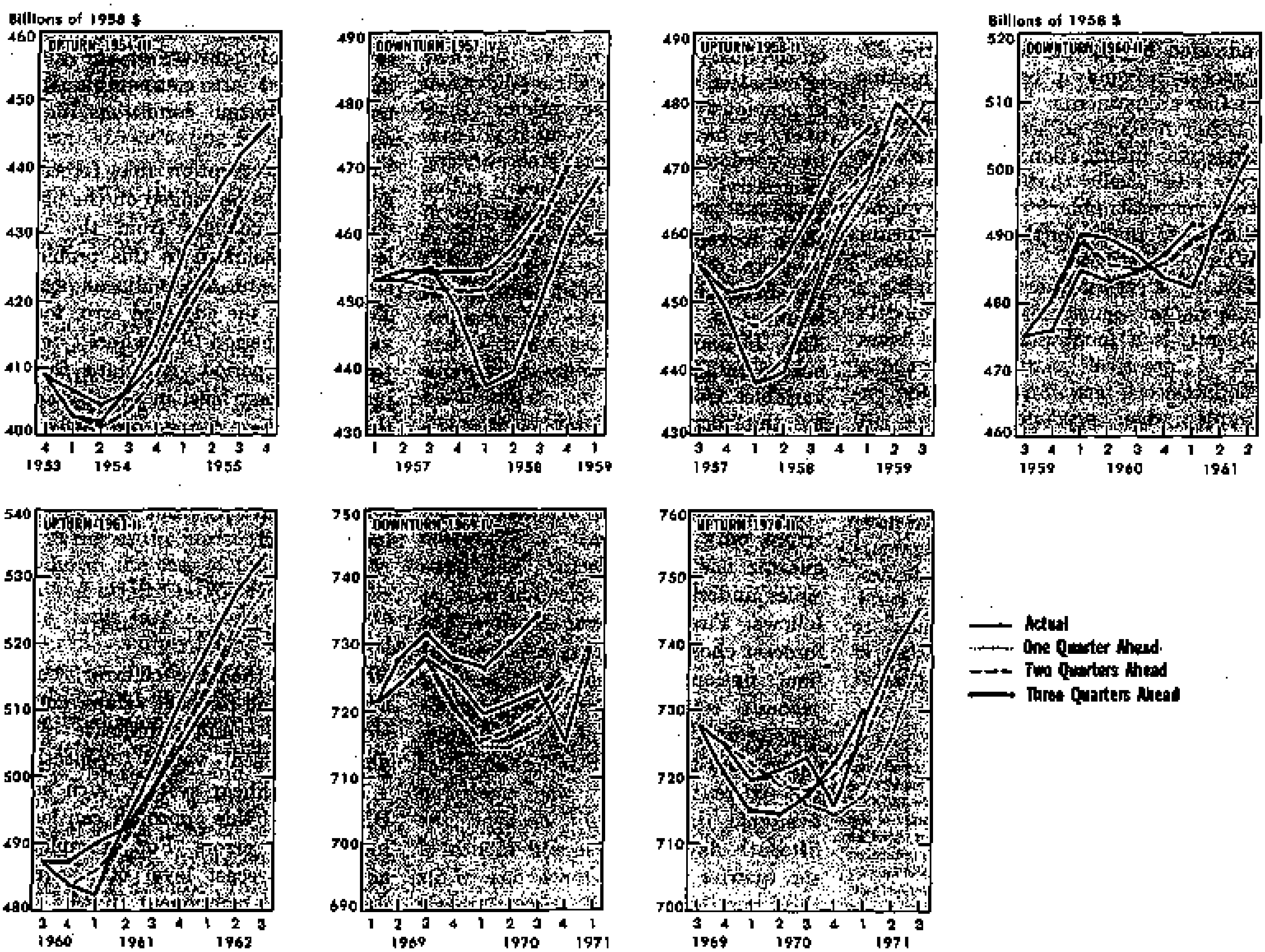
15. This quarter is set as the cutoff point primarily because the new price and wage policy, beginning with the freeze imposed on August 15, 1971, introduced an important structural change that was not anticipated.

sequentially for the desired period. To most practitioners of econometric forecasting, however, such a procedure appears inadequate. A model can be a powerful aid to forecasting, but it should not be a straightjacket. Most experienced model forecasters exercise considerable control over their model's output by departing from mechanical procedures. Such departures are based on internal information (nature of past equation residuals), external information (knowledge about the economy that is not incorporated in the model structure or is inconsistent with preliminary model outputs), or judgmental restrictions on the outputs.

Judgmental elements (other than projections of exogenous variables) are usually introduced by using constant adjustments that do not rest on a formula. Such adjustments are mostly made prior to running a forecast. The following are specific reasons—based on internal or external information—why such prior adjustments are made: First, the recent pattern of residuals may not suggest the "decay" process given by the formula. For instance, one would obviously not want to use the formula employed in the post-sample predictions if recent residuals show a clear trend; rather, one would tend to continue the trend in the adjustments. Second, there may be special factors which explain the most recent residuals, but which are not relevant to the forecast period or are relevant in a special way. Third, the forecaster may know about an impending circumstance, either as a certainty or as a substantial possibility, which calls for

CHART 10

Cyclical Turning Points in Real GNP



Seasonally Adjusted at Annual Rates

U.S. Department of Commerce, Bureau of Economic Analysis

78-8-10

special adjustment; or he may regard certain factors that are not incorporated into the model structure as relevant during the period of forecast. Examples are: legislation having some future effect; a strike threat which is expected to lead to certain kinds of anticipatory behavior; the direct effects of a strike on production; the effect on consumption of a large and prolonged decline in stock prices.

There is one kind of external information in response to which adjustments are (normally) made only after an initial run is obtained. Since the first quarter covered by each forecast is usually well in progress at the time of making a forecast, there is partial information already available on developments in that quarter. Thus, if model outputs obviously conflict with what is indicated by the partial data, the differences become a basis for making further adjustments.

After making adjustments based on objective information, the forecaster may still decide that certain outputs are "unreasonable," and he might, therefore, make further adjustments. What constitutes "reasonableness," that is, a valid basis for imposing a judgmental constraint on a model's output, is a moot philosophical issue. As a practical matter, judgmental modifications are introduced, for example, when predicted changes are strikingly large or small compared to past changes or when certain ratios are well outside their historical range. It should be emphasized that we have tried to refrain from making judgments about global or summary magnitudes, such as total GNP or the price level, but rather have confined them to specific items, such as components of GNP, and to notions of "consistency" among variables.[15]

At times certain behavioral relationships break down to such an extent that it becomes easier to substitute extraneous estimates of certain variables for endogenously determined values than to try to modify the results by constant adjustments. This has, for instance, sometimes been the case with housing starts. Another example: Equations based upon past market determination of wages and prices become inappropriate if price-wage policies such as those initiated in August 1971 are effective in modifying price and wage behavior.

Selection of forecasts and computation of errors

We have frequently made more than one forecast during a quarter because of significant data revisions, new data on the quarter in progress, or other new information calling for the modification of previously made assumptions. Moreover, we have often presented alternative versions of a forecast incorporating alternative assumptions about policy decisions or about the occurrence of some exogenous future event, such as a strike, or alternative time paths of certain exogenous variables about which there was considerable uncertainty. Also, in some instances where it appeared unlikely that certain endogenous variables would turn out as predicted by the model, alternative forecasts, making these variables exogenous, were run.

For the purpose of analyzing forecast errors, only one forecast made in each quarter has been selected. The principles underlying the selection were as follows: First, wherever possible, the forecast chosen was one made after the final national income and product account estimates for the previous quarter had been completed, but before substantial two-month information for the current quarter was becoming available, i.e., roughly between the fifteenth day of the second month of the quarter and the tenth day of its third month. Second, if a given forecast had versions with differing degrees of endogeneity, that with the maximum endogeneity was selected. Third, where more than one fiscal policy or strike variant was available, the variant whose assumption most closely approximated the actual subsequent event was selected.

The periodic revision of data creates a problem for the measurement of forecasting error, but one which can be essentially overcome. It is assumed that revised data are more accurate than preliminary data and hence are a more appropriate basis against which to evaluate forecasts. Since the initial conditions and the data being forecast tend to be revised in the same direction, some kind of adjustment of the forecast for revisions in the initial conditions is warranted. (When a forecast horizon extends over the time of the annual (July) revisions of the national income and product accounts, unrevised data for the final quarters of the forecast horizon that are comparable to the unrevised initial conditions do not even exist.)

On the assumption that forecasting accuracy is to be judged on the basis of cumulative changes from the initial levels, the solution to this problem is straightforward for a linear system: To compute the adjusted error for any variable in period $t+i$, where t denotes the initial period, calculate the adjusted forecast level in $t+i$ by adding the cumulative change originally forecast from t to $t+i$ to the revised initial level; the revised actual level in $t+i$ is then subtracted from the adjusted forecast level:

$$e_{t+i}=F^r_{t+i}-A^r_{t+i}$$
$$=(A^r_t-F^u_{t+i}-A^u_t)-A^r_{t+i},$$

where e_{t+i} is the adjusted error, A and F are actual and forecast values, and the superscripts r and u indicate revised and unrevised values respectively.

In a nonlinear system, such as the BEA model, this approach can lead to inconsistent errors. For example, the adjusted forecast level of a current-dollar GNP component in $t+i$ is, in general, not precisely the same when current-dollar values are used directly as it is when the separately forecast

15. This assertion contradicts the following conclusion based on a study of BEA(OBE) and Wharton model forecasts by Yoel Haitovsky and George Treyz: "We find that there is reason to believe that in the first quarter of the forecast, both the values chosen for the exogenous variables and the discretionary constant adjustments were influenced by interaction between the forecaster and the model forecast and that this interaction improved . . . first quarter forecasts [of GNP]" ("Forecasts with Quarterly Macroeconometric Models, Equation Adjustments, and Benchmark Predictions," *Review of Economics and Statistics*, August 1972, p. 320). This statement suggests that the forecasters systematically adjust constants in such a way as to offset the effect of errors in the exogenous variables on the forecasts of GNP.

real GNP components and implicit price deflators are used. However, over relatively short forecasting horizons the inconsistency is not likely to be serious. Thus, the approach described above was used to calculate adjusted error, following the convention that the adjustment is applied directly in each case to the variable that is the subject of error measurement.

The plan for the remainder of this section is as follows: First, RMSE and bias statistics for *ex-ante* (i.e., actual) forecasts are presented. Errors in *ex-ante* forecasts are then compared with errors in post-sample predictions. Next, errors in *ex-ante* forecasts are compared with errors in corresponding "*ex-post*" forecasts (adjusted *ex-ante* forecasts with actual values of exogenous variables substituted for the forecast values) in order to determine the effect of errors in forecasting exogenous variables. Finally, after a brief discussion of the problems of evaluating forecasting performance, comparisons are made between errors in the *ex-ante* forecasts and those based on extrapolation using autoregressions as "naive" benchmarks.

Size of errors in ex-ante forecasts

Table 7 shows root mean square errors for major variables from *ex-ante* forecasts with horizons up to six quarters. As is indicated in the first line of the table, the number of forecasts is not constant for the various horizons, but diminishes as the horizon lengthens.

The RMSE's show a general tendency to rise as the horizon lengthens, as they did in the case of the predictions. (A falling off in the RMSE's for many of the items in the table in the sixth quarter may simply reflect the fact that only five of the 22 forecasts are represented for that horizon.) As in the case of the predictions, there is a tendency for offsetting errors among GNP components. Moreover—and this is a new feature—there is a very strong offsetting of errors between real GNP and the price level (represented by the private GNP deflator), resulting in only moderate errors in current-dollar GNP. As shown below, forecast errors in these variables show bias in opposite directions.

Errors in personal consumption expenditures and in business inventory investment are heavy contributors to errors in current-dollar GNP forecast one quarter ahead. Errors in forecasting consumer purchases of autos and parts (shown only in 1958 dollars in the table) account for a major portion of the errors in forecasting total personal consumption expenditures, despite the fact that autos and parts purchases are only about one-tenth of the total. Errors in fixed nonresidential investment and imports become prominent by the fifth quarter.

The RMSE's for each of the two main exogenous components of GNP—exports and government purchases—are also substantial, but the RMSE for the sum of the two components is considerably less than the sum of the separate RMSE's, reflecting substantial offsetting of errors. The size of errors in the sum of government purchases relative to that of errors in total GNP diminishes as the forecast horizon lengthens.

Forecast errors in personal income increasingly dominate those in corporate profits as the forecast horizon lengthens. In percentage terms (not shown), however, RMSE's for corporate profits are uniformly larger.

Bias

As noted above, errors in forecasts of current-dollar GNP are moderated by offsetting errors in forecasts of prices and real GNP. This offsetting reflects a positive bias in forecasts of real GNP and a substantial negative bias in

Table 7.—Root Mean Square Errors of Selected Variables: *Ex-ante* Forecasts (1966–I—1971–II)

	Forecast horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Number of observations	22	21	20	16	11	5
Billions of dollars:						
Gross national product	2.94	5.67	10.60	15.27	16.96	16.48
Change in GNP	2.04	6.60	6.60	7.60	7.42	6.94
Personal consumption expenditures	1.93	6.60	9.32	11.67	12.09	11.69
Fixed nonresidential investment	1.67	2.84	3.18	4.34	5.68	5.02
Residential structures	1.14	1.79	2.83	3.38	3.24	3.64
Change in business inventories	3.78	4.15	4.19	4.77	3.88	2.00
Exports of goods and services*	1.70	2.14	2.29	2.52	3.40	3.19
Imports of goods and services	1.84	2.52	3.12	3.84	5.19	5.98
Government purchases of goods and services*	1.86	2.83	3.63	5.05	4.46	4.94
Exports plus government purchases*	2.24	3.13	3.95	4.87	3.65	4.19
Personal income	2.69	6.02	9.19	13.16	16.48	18.78
Corporate profits and inventory valuation adjustment	2.74	4.39	5.16	6.47	10.26	8.56
Billions of 1958 dollars:						
Gross national product	2.89	5.61	8.78	11.94	15.84	17.27
Personal consumption expenditures	2.54	4.29	6.01	7.35	7.86	9.60
Automobile and parts	1.60	2.58	2.87	3.50	3.80	4.67
Nonauto durables	.70	.85	1.26	1.73	1.71	1.76
Nondurables	2.14	1.91	2.78	3.62	2.78	3.28
Nonhousing services	.75	1.22	1.62	1.89	2.63	3.07
Housing services	.29	.37	.58	.63	.64	.62
Fixed nonresidential investment	2.48	2.85	2.46	3.45	3.87	3.64
Residential structures	.77	1.19	1.86	2.46	2.06	2.74
Change in business inventories	3.42	3.81	3.64	4.43	3.48	1.81
Merchandise imports**	1.87	1.76	2.19	2.51	2.70	2.65
Services imports, nondefense**	.46	.70	.83	1.11	1.38	1.20
Miscellaneous variables:						
Implicit price deflator, private GNP (1958=100)	.44	.75	1.16	1.63	2.61	2.47
Wages per private employee (dollars per year)	41.2	66.6	87.2	120.6	128.5	113.1
Index of private output per manhour (1958=100)**	.87	1.84	1.92	2.68	3.29	3.92
Civilian labor force (millions)	.28	.45	.61	.69	1.08	1.11
Employed	.29	.49	.66	.78	1.04	.91
Unemployment rate (percent)	.13	.37	.38	.58	.74	.84
Average yield on 4-6 months commercial paper (percent)	.47	.93	1.32	1.85	2.34	2.46
Average yield, corporate bonds (percent)	.23	.52	.82	1.09	1.41	1.56
	Root Mean Square Percentage Errors					
Gross national product	.60	.80	.95	1.13	1.10	1.10
GNP in 1958 dollars	.60	.82	1.05	1.23	1.41	1.46
Implicit deflator, private GNP	.55	.70	.89	1.10	1.34	1.43

*Exogenous.
**Based on forecasts since April 1967 only.

Table 8.—Bias in *Ex-ante* Forecasts for Major Items

[Average errors and standard deviations are in $billions, except as otherwise indicated]

	Forecast horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Gross national product:						
$(\bar{P}-\bar{X})$[a]	−0.90	−1.60	−1.77	−3.87	−7.02	−2.48
S_{P-X}[b]	.93	2.10	3.66	4.98	5.15	5.00
$[(\bar{P}-\bar{X})^2/MSE]$[c]	.085	.058	.028	.041	.171	.024
GNP in 1958 dollars:						
$\bar{P}-\bar{X}$	1.43	2.88	4.57	6.24	6.89	18.46
S_{P-X}	.84	1.69	2.49	3.40	4.27	3.89
$[(\bar{P}-\bar{X})^2/MSE]$	.235	.183	.274	.270	.345	.611
Implicit deflator, private GNP:						
$\bar{P}-\bar{X}$	−.32	−.54	−.90	−1.24	−2.28	−2.64
S_{P-X}	.20	.17	.22	.36	.20	.35
$[(\bar{P}-\bar{X})^2/MSE]$	.485	.492	.634	.548	.984	.866
Personal consumption expenditures:						
$\bar{P}-\bar{X}$	−.49	−.08	−.30	−1.43	−7.63	−6.69
S_{P-X}	1.29	2.27	3.10	3.63	3.15	3.88
$[(\bar{P}-\bar{X})^2/MSE]$	.016	.000	.001	.012	.888	.229
Fixed nonresidential investment:						
$\bar{P}-\bar{X}$	−.16	−.12	−.35	−1.76	−2.65	−2.84
S_{P-X}	.55	.86	1.05	1.32	1.67	1.49
$[(\bar{P}-\bar{X})^2/MSE]$	.069	.002	.012	.164	.218	.264
Residential structures:						
$\bar{P}-\bar{X}$	−.39	−.71	−.40	−.81	−.46	−.64
S_{P-X}	.98	.54	.80	1.09	1.07	.86
$[(\bar{P}-\bar{X})^2/MSE]$	.117	.189	.190	.067	.020	.069
Change in business inventories:						
$\bar{P}-\bar{X}$	−.51	−.77	−.08	.09	.47	1.16
S_{P-X}	1.25	1.38	1.40	1.69	1.29	.54
$[(\bar{P}-\bar{X})^2/MSE]$	.018	.034	.000	.000	.016	.394
Exports & government purchases*:						
$\bar{P}-\bar{X}$	−.44	−1.40	−2.20	−2.96	−1.65	−.14
S_{P-X}	.72	.96	1.11	1.29	1.10	1.40
$[(\bar{P}-\bar{X})^2/MSE]$	.038	.200	.306	.367	.180	.001
Imports of goods and services:						
$\bar{P}-\bar{X}$	−.66	−1.56	−2.20	−2.78	−4.86	−5.96
S_{P-X}	.84	.86	.71	.82	.41	.18
$[(\bar{P}-\bar{X})^2/MSE]$	.213	.383	.497	.524	.877	.992
Personal income:						
$\bar{P}-\bar{X}$	−1.46	−3.05	−4.45	−6.05	−14.11	−13.60
S_{P-X}	.75	1.73	2.69	3.47	3.95	4.23
$[(\bar{P}-\bar{X})^2/MSE]$	.204	.257	.248	.274	.586	.555
Corporate profits and inventory valuation adjustment:						
$\bar{P}-\bar{X}$	.37	1.71	2.79	2.95	5.84	5.62
S_{P-X}	.89	1.38	1.45	1.71	2.02	2.15
$[(\bar{P}-\bar{X})^2/MSE]$	.043	.408	.290	.373	.371	.431
Unemployment rate (percent):						
$\bar{P}-\bar{X}$	−.02	−.00	−.06	−.12	−.17	−.42
S_{P-X}	.06	.12	.12	.17	.24	.24
$[(\bar{P}-\bar{X})^2/MSE]$	.028	.000	.024	.081	.054	.218

a Average prediction error.
b Standard deviation of average prediction error.
c Bias proportion (square of average error as a proportion of mean square error).
*Exogenous.

forecasts of the deflator, as can be seen in table 8 (comparable to tables 4 and 5). The bias in the real GNP forecasts becomes significant at the 5 percent level after the fourth quarter. For the deflator, bias is significant for all quarters.

The positive bias in the real GNP forecast errors reflects a positive, though not significant, bias in the forecast errors for real personal consumption expenditures and a strong negative bias in forecast errors for real merchandise imports (not shown in the table). In current dollars, there are large and significant negative errors in total imports, while the sum of the two main exogenous variables (exports and government purchases) also has a negative mean error for all quarters with bias significant in the third and fourth quarters. On the income side, a strong negative bias in the personal income errors is partly offset by a positive, but marginally significant, bias in corporate profits. Errors in the unemployment rate show no significant bias.

The source of bias in forecasting errors is not clear. However, through a systematic decomposition of errors by sources of error—which we intend to undertake in the near future—we can identify equations that are critical in producing bias and make adjustments that tend to eliminate it. Reduction of bias, of course, serves to reduce the size of errors.

Ex-ante *forecasts versus post-sample predictions*

The *ex-ante* forecasts reviewed in this article differ from the post-sample predictions in the following ways: (1) The forecasts cover a longer time period; (2) they incorporate a mixture of model versions and a somewhat varying degree of exogeneity; (3) they use judgmental projections of exogenous variables rather than actual values; (4) they embody many judgmental (in place of mechanical) constant adjustments; (5) they use unrevised rather than revised data as initial conditions.

Because of the multiplicity of differences, it is difficult to compare error statistics from the two sets of runs in a meaningful way. It is also hazardous to generalize from them because of the extreme shortness of the period covered by the post-sample predictions. Having both sets of results, however, makes a comparison between them irresistible, and we shall try to draw such tentative inferences as we can.

Comparison of table 7 with table 2 reveals that for both current-dollar GNP and real GNP, RMSE's of the *ex-ante* forecasts are in almost all cases smaller than the RMSE's of the post-sample predictions. For current-dollar GNP, the comparative sizes of the errors reflects primarily the fact that offsetting biases between price and output errors occur in the forecasts, but not in the predictions.

RMSE's for components of real GNP are also generally smaller in the forecasts. However, the relative composition of errors is broadly similar in the two cases; in the case of the forecasts, errors in personal consumption expenditures dominate those in total GNP to a somewhat greater extent.

Errors in the private GNP deflator, besides being negatively biased, are substantially larger in the forecasts than in the predictions. For horizons up to five quarters, the root mean square percentage errors are almost twice as large.

Forecast errors in the unemployment rate are smaller than those in the predictions, a result that is consistent with the smaller errors in real GNP. Errors in forecasting both short- and long-term interest rates are generally much larger than in the predictions. This stems in large part from failure to forecast correctly changes in the discount rate, which is exogenous and which has a strong impact on the short-term rate and an indirect and weaker short-run impact on the long-term rate.

In order to remove, at least for the aggregative output and price variables, differences between forecasts and post-sample predictions that are due to the difference in the time period covered, RMSE's for these variables were also computed for the subset of forecasts covering the same period (1969–I to 1971–II) used for the predictions. Table 9 shows these RMSE's for current-dollar GNP, real GNP, and the implicit private GNP deflator. The same general pattern emerges as in the comparison of the full set of forecasts with the predictions. For current-dollar GNP, RMSE's in the subset of forecasts are smaller than in the the full set, averaging only about one-third as large as in the predictions. After the first quarter, real GNP errors are larger in the subset of forecasts than in the full set, and are, therefore, closer to the prediction errors; in the fifth and sixth quarters they are greater. Price errors are generally somewhat greater in the subset of forecasts than in the full set, and thus show an even wider margin over the price errors in the post-sample predictions. There is evidently more offsetting of errors between real output and the price level in the 1969–71 subset of forecasts than in the full set.

Table 9.—Root Mean Square Errors of Selected Variables: *Ex-ante* Forecasts Versus Post Sample Predictions (1969–I—1971–II)

	Forecast or prediction horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Gross national product (billions of dollars):						
Ex-ante forecasts	2.07	4.67	5.48	6.61	8.88	9.89
Post-sample predictions	6.01	11.01	18.42	23.63	28.06	30.5
Gross national product (billions of 1958 dollars):						
Ex-ante forecasts	2.37	5.65	9.22	13.26	16.40	20.84
Post-sample predictions	4.56	6.11	13.29	16.06	17.78	17.16
Implicit deflator, private GNP (1958=100):						
Ex-ante forecasts	.51	.67	1.32	1.55	2.45	3.16
Post-sample predictions	.27	.44	.41	.70	.95	1.31

The relatively poor performance of the forecasts of the price level, as compared with the predictions, is largely explained by bigger errors in productivity (output per man-hour) in the forecasts than in the predictions (judging from comparisons based on the full set of forecasts). Another possible explanatory factor is that a recently introduced method of solving for the price level, which has been shown to reduce errors,[17] was used in all of the predictions, but in only the last few of the forecasts. Errors in the private sector wage rate—the other element of unit labor cost, which is the main determinant of the price level—are somewhat smaller in the forecasts than in the predictions.

The reason why errors in real GNP are smaller in the *ex-ante* forecasts than in the post-sample predictions is less obvious. The above comparisons do not indicate offsetting of larger errors among components of real GNP; nor do the results of the next section show substantial offsetting between errors in exogenous variables and model prediction errors. A likely explanation—though a tentative one—is that judgmental adjustments have contributed to forecasting accuracy, at least for the first few forecast quarters.[18]

Ex-post *versus* ex-ante *forecasts*

The main judgmental element in econometric forecasting is the projection of exogenous variables. It is thus of interest to determine whether errors in these projections have worsened the accuracy of the GNP forecasts, and if so, to what extent. This can be ascertained by repeating forecasts made in the past with all inputs other than the exogenous variables kept intact; for the latter, the actual values are used. This type of repetition of past *ex-ante* forecasts is called *ex-post* forecasting.

What cannot be done, unfortunately, is to update the judgmental elements in the constant adjustments—that is, departures from mechanical adjustments—in accordance with data revisions and

17. Albert A. Hirsch, "Price Simulations with the OBE Econometric Model," in *Econometrics of Price Behavior*, edited by Otto Eckstein, Board of Governors of the Federal Reserve System, 1972.

18. This inference is corroborated by evidence in the study by Haitovsky and Treyz (*op. cit.*). For both the BEA and Wharton Model forecasts that are analyzed, the forecasts of current-dollar GNP, real GNP, and the unemployment rate are generally poorer when the judgmental constant adjustments are replaced by mechanical adjustments (but the *ex-ante* exogenous variables are used). We are undertaking a more intensive investigation of the role of different kinds of judgment in forecasting and we will report the results when more observations become available.

Table 10.—Root Mean Square Errors: *Ex ante* versus *Ex-post* Forecasts (1967–II—1971–II)

	Forecast horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Number of observations	17	16	15	13	9	4
Gross national product (billions of dollars):						
Ex-ante	2.45	7.06	10.98	14.95	18.48	17.97
Ex-post	3.82	9.43	14.58	19.11	22.04	18.41
Gross national product (billions of 1958 dollars):						
Ex-ante	2.59	5.92	9.11	12.29	17.39	18.74
Ex-post	2.61	6.61	10.24	13.42	14.38	14.69
Implicit deflator, private GNP (1958=100):						
Ex-ante	.46	.78	1.28	1.78	2.57	3.12
Ex-post	.38	.76	1.20	1.79	2.79	3.79
Unemployment rate (percent):						
Ex-ante	.14	.29	.42	.59	.81	.94
Ex-post	.16	.35	.56	.72	.86	.85

ex-post knowledge of special factors. The problem is that this would have to be done in a way that is not prejudiced by the actual outcome of the data being forecast. Putting it in another way, so far as the constant adjustments are concerned, we cannot disentangle the uncertainty of the forecaster's judgment from the "uncertainty" (i.e., purely stochastic elements) inherent in the model. Thus, in our effort to remove errors of judgment, we are limited to the removal of errors in projecting exogenous variables.

Reference has already been made to the study by Haitovsky and Treyz of *ex-ante* and *ex-post* forecast errors in the BEA and Wharton models (footnote 14); the period covered for the BEA model in that study is 1967–III through 1969–III. It has not been possible to replicate forecasts made prior to 1967–III, but results of the forecasts covering 1969–IV through 1971–II have been added. Thus, *ex-post* forecasts can be examined for all but the first five of the full set of *ex-ante* forecasts.

In the *ex-post* forecasts, values of the exogenous variables were determined by linking changes in them, as measured by data now available, to the initial levels used in the *ex-ante* forecasts. Adjusted errors in the endogenous variables were then computed in the same way as described for the *ex-ante* forecasts. Table 10 compares root mean square errors in current-dollar GNP, real GNP, the private GNP deflator, and the unemployment rate for this set of *ex-post* forecasts with those of the corresponding *ex-ante* forecasts.

The results are mixed. Forecasts of current-dollar GNP are uniformly and substantially poorer *ex-post*. We have not compiled all the necessary data to determine precisely why this rather surprising result occurs. As in the case of the analogous superiority of *ex-ante* forecasts over post-sample predictions of current-dollar GNP, it must result from offsetting errors. In this instance, however, the offsetting occurs between errors in the exogenous variables and the errors in the model equations. For real GNP, RMSE's of *ex-post* forecasts are only slightly worse in the first four quarters and are substantially better in the fifth and sixth quarters. The RMSE's for the *ex-post* and the *ex-ante* forecasts of the private GNP deflator are almost identical until the sixth quarter. *Ex-post* forecast errors in the unemployment rate are slightly larger until the sixth quarter, primarily reflecting the larger *ex-post* errors in real output.

While there are no formal hypothesis tests to ascertain whether the differences between *ex-ante* and *ex-post* forecast errors are statistically significant, the small differences for variables other than current-dollar GNP and the small number of degrees of freedom strongly suggest that the differences are not significant.

Further evaluation of forecasting performance

Comparisons have been made between *ex-ante* forecast errors and prediction errors and between *ex-ante* and *ex-post* forecast errors. However, neither set of comparisons answers the question: Is an econometric model (and the BEA model in particular) a useful device for forecasting?

This question cannot be answered unambiguously for at least the following reasons: (1) There is no single criterion of forecasting quality; (2) what constitutes a sufficiently accurate forecast depends on how the results are to be used; (3) as previously noted, one cannot fully separate the use of a model in forecasting from the judgment of the forecaster; hence, various tests that attempt to assess the contribution of the model to forecasting are not wholly satisfactory; (4) the continuing process of model development and improvement and the changing structure of the economy imply that the past record of a model is not a clear guide to its future performance.

In regard to the first point, it must be borne in mind that a forecast does not yield a single magnitude, but a whole vector of outputs (and for multiperiod forecasts, a set of such vectors or matrix of outputs). A forecaster or forecasting system can, for instance, produce an excellent record in terms of current-dollar GNP, but a poor one in terms of composition of final demand, prices, profits, etc. Again, a set of quarterly forecasts might include good one- and two-quarter projections, but poor ones beyond that; or it may provide good year-ahead forecasts, but give misleading quarterly patterns and fail to indicate cyclical turning points. There is no objective basis for weighting these various elements to arrive at an overall rating of forecasting quality.

One may compare model forecasts with forecasts obtained by other techniques, but here also there are problems. For instance, the record of an econometric forecaster, or of a group of econometric forecasters, can be compared with that of a selected group of judgmental forecasters over the same period. However, it is not likely that a reputable judgmental forecaster exists who can claim to be uninfluenced by econometric forecasts (and probably also vice versa) and there is an obvious problem in deciding which econometric and judgmental forecasts are to be compared. Another obstacle to meaningful comparison is that judgmental and econometric forecasters do not project a common set of values for the exogenous variables of the models.[19]

Standardized comparisons of the forecasting ability of different econometric models are also hard to obtain. The difficulty in sorting out the forecaster's judgmental errors from the performance of the model is a particular problem here.[20]

Comparison of ex-ante forecasts with "naive" benchmarks

In view of the problems of finding an absolute standard for evaluating forecasts and in comparing model forecasts with judgmental forecasts or one model's forecasts with those of another, it is worthwhile to make still another kind of comparison, namely, between *ex-ante* forecasts and so-called "naive" extrapolations. The latter approach is one that makes no use of economic knowledge other than past data on the variables in question and hence is merely a mechanical device for projecting the data. Such a benchmark can represent a sort of floor below which, it is hoped, forecasting performance based on the non-naive method will not sink.

A whole spectrum of naive benchmarks, varying in degree of complexity, is available. The simplest one is a projection of no change from the previous period. Since the economy is generally upward trending, it is clear that all serious forecasting systems would win over this benchmark. A benchmark that provides a somewhat harder test is extrapolation of the same change as in the previous period. A more complex benchmark, which involves statistical inference applied to past economic data but which again fails to incorporate hypotheses of economic casuality, is an autoregressive equation, that is, one in which the variable in question depends on its own lagged values:

$$Y_t = a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_n Y_{t-n}.$$

Table 11 shows RMSE's for extrapolations of major variables one to six quarters ahead made using "second-order" autoregressive equations, that is, equations having two lagged values of the dependent variable. Inclusion of the second as well as the first lagged value results in difference equations which may yield cyclical movements as well as growth. As in the predictions and forecasts made with the model, the autoregressive extrapolations were generated dynamically; that is, extrapolated rather than actual values of the lagged dependent variables were used where needed as inputs.

Extrapolations with the autoregressive equations were made for the same

Table 11.—Root Mean Square Errors for Selected Items: *Ex ante* Forecasts versus Extrapolations using Second-order Autoregressive Equations

	Forecast horizon (quarters)					
	1Q	2Q	3Q	4Q	5Q	6Q
Billions of dollars:						
Gross national product:						
Ex-ante forecasts	2.94	6.67	10.60	15.27	16.95	15.48
Autoregressive	4.53	7.44	10.62	14.28	17.68	19.61
Personal income:						
Ex-ante forecasts	2.69	6.02	9.19	13.16	13.43	16.78
Autoregressive	2.41	3.74	5.66	6.74	7.80	8.51
Corporate profits and inventory valuation adjustment:						
Ex-ante forecasts	2.74	4.29	5.16	6.47	10.26	8.56
Autoregressive	2.58	3.65	4.90	6.44	7.99	8.88
Billions of 1958 dollars:						
Gross national product:						
Ex-ante forecasts	2.89	5.61	8.79	11.94	15.94	17.27
Autoregressive	4.35	7.29	11.41	16.75	22.62	28.27
Personal consumption expenditures:						
Ex-ante forecasts	2.54	4.29	6.01	7.35	7.38	9.50
Autoregressive	3.79	5.25	7.21	10.23	13.32	16.02
Automobiles and parts:						
Ex-ante forecasts	1.60	2.69	2.87	3.66	3.50	4.67
Autoregressive	2.39	2.14	2.06	2.74	2.70	2.45
Fixed nonresidential investment:						
Ex-ante forecasts	2.48	2.36	2.46	2.46	2.87	3.84
Autoregressive	1.37	2.40	3.99	6.06	8.40	10.69
Residential structures:						
Ex-ante forecasts	.77	1.19	1.86	2.16	2.06	2.24
Autoregressive	.78	1.41	1.98	2.31	2.45	2.46
Change in business inventories:						
Ex-ante forecasts	3.49	3.81	3.64	4.43	3.48	1.61
Autoregressive	4.41	5.19	5.67	6.88	7.05	7.24
Merchandise imports:						
Ex-ante forecasts*	1.67	1.78	2.19	2.51	2.70	2.85
Autoregressive	1.19	1.61	1.47	1.60	1.52	1.88
Miscellaneous items:						
Implicit price deflator, private GNP (1958=100):						
Ex-ante forecasts	.44	.75	1.16	1.63	2.64	2.47
Autoregressive	.31	.66	1.01	1.37	1.86	2.49
Unemployment rate (percent):						
Ex-ante forecasts	.13	.37	.38	.63	.74	.84
Autoregressive	.23	.50	.83	1.14	1.38	1.56
Average yield on 4-6 months commercial paper (percent):						
Ex-ante forecasts	.47	.98	1.32	1.85	2.34	2.46
Autoregressive	.38	.62	1.15	1.38	1.57	1.09
Average yield, corporate bonds (percent):						
Ex-ante forecasts	.23	.52	.85	1.08	1.41	1.58
Autoregressive	.78	1.59	2.20	2.76	2.81	.87

*Based on forecasts since April 1967 only.

19. Notwithstanding these difficulties, a very tentative comparison between model and judgmental forecasts has been made by Victor Zarnowitz, using the BEA and Wharton model results obtained by Haitovsky and Treyz. According to Zarnowitz, models have "a slight edge" on judgmental forecasts; "Forecasting Economic Conditions: Record and Prospect," *The Business Cycle Today*, edited by Victor Zarnowitz, National Bureau of Economic Research, 1972, especially pp. 222-27.

20. A cooperative attempt to make intermodel comparisons of predictive ability and other properties, by imposing procedures as uniform as possible, has been undertaken by various model builders under the sponsorship of the National Bureau of Economic Research and the National Science Foundation.

Preliminary comparisons of sample period and post-sample period predictions of current-dollar GNP and real GNP for all models, including the BEA model, have been published in Lawrence R. Klein and Gary Fromm, "A Comparison of Eleven Econometric Models of the United States," *American Economic Review*, May 1973. (Comparisons of *ex-ante* forecasts have not been made by this group.) The results indicate relatively little variation across models in RMSE's within the sample period. There is more variation in the post-sample comparisons; however, these comparisons are hampered by lack of perfect uniformity in the time period covered, shortness of the post-sample period in most cases, and adjustment procedures.

For the period 1967-III—1969-III, comparisons of *ex-ante* and *ex-post* forecast results for key variables from the BEA and Wharton models using actual and alternative constant adjustment procedures have also been published (Haitovsky and Treyz, *op. cit.*).

period as that for which the model forecasts are presented, i.e., 1966—1971–II. However, two sets of parameter estimates were made for each autoregressive equation: The first set was estimated over the period 1953–II—1966–IV and used to extrapolate in 1967 and 1968; the second set was estimated through 1968 and used to extrapolate in 1969—1971–II. This reestimation of the autoregressive equation parameters corresponds to the timing of major reestimations of the model.

For current-dollar GNP, the model *ex-ante* forecasts are, on balance, superior to the autoregressive extrapolations, although a reversal occurs in the third and fourth quarters. The model forecasts of real GNP are distinctly superior to autoregressive extrapolations throughout, while the forecasts of the private GNP deflator are inferior to the extrapolations until the sixth quarter. Perhaps the autoregressive price equation captures the role of price expectations to a substantial degree.

Among components of real GNP, model forecasts are better than autoregressive extrapolations of total consumption expenditures, fixed nonresidential investment after the first quarter, business inventory investment, and, to a slight extent, residential construction outlays. However, model forecasts are weaker for consumer purchases of autos and parts after the first quarter and merchandise imports in all quarters.

Beyond one quarter, the model forecasts of personal income yield substantially larger errors than the autoregressive extrapolations. For profits, the two sets of errors are closer, although the model forecasts are slightly poorer. The model forecasts are superior for the unemployment rate and the long-term interest rate, but not for the short-term interest rate.

The mixed performance of the *ex-ante* forecasts relative to extrapolation based on autoregression is somewhat disappointing. However, this does not suggest that we should abandon econometric models in favor of autoregressive or other empirical extrapolation techniques. Rather, it points up the need for further improvement in model structure or, in statistical techniques of estimation or prediction. Even if they would consistently yield more accurate forecasts, purely empirical techniques could not deal with the impact of assumed alternative policy decisions and other contingencies. Only a model can do this.

Table 12.—Errors in Forecast of August 8, 1968

[Predicted minus actual]

	1968		1969			
	III	IV	I	II	III	IV
Billions of dollars (except as indicated):						
Gross national product	−7.7	−14.6	−23.0	−33.8	−37.8	−31.4
Personal consumption expenditures	−8.4	−11.0	−17.5	−22.5	−22.2	−24.3
Automobiles and parts	−2.2	−3.7	−3.4	−3.4	−5.5	−6.8
Other durable goods	−1.1	−1.3	−2.4	−3.3	−1.1	−.9
Nondurable goods	−2.5	−3.8	−6.5	−9.2	−8.9	−9.1
Services	−1.7	−2.3	−3.3	−4.7	−6.7	−5.3
Fixed nonresidential investment	.4	−1.2	−.1	−3.9	−5.7	−9.8
Residential structures	.1	−2.0	−4.3	−4.8	−3.5	−.9
Change in business inventories	−1.2	−2.5	−4.2	−4.2	−7.2	−1.7
Net exports	1.2	3.5	3.8	4.0	2.9	2.9
Exports	−1.6	1.6	5.2	−2.9	−3.4	−3.7
Imports	−2.8	−2.0	1.6	−7.0	−6.3	−6.6
Government purchases of goods and services	.2	−1.5	.6	1.3	1.9	2.2
Federal	.2	−.5	1.5	2.8	2.9	3.1
State and local	.0	−1.0	−.9	−1.5	−1.0	−.9
Personal income	−3.9	−11.0	−19.4	−27.6	−32.4	−36.4
Disposable personal income	−2.4	−7.7	−10.4	−16.6	−24.7	−29.0
Corporate profits and inventory valuation adjustment	−2.6	−1.5	−2.2	−2.7	.5	7.5
Gross national product in 1958 dollars	−5.0	−7.5	−12.8	−14.2	−13.8	−8.0
Implicit price deflator, private GNP (1958=100)	−.2	−.7	−1.2	−1.8	−2.6	−2.4
Wages per private employee (thousands of dollars per year)	−.06	−.15	−.16	−.24	−.34	−.22

Summary of Major Findings

(1) Both prediction and forecasting errors tend to grow in size as the prediction horizon is extended. This tendency is greatly subdued in predictions and forecasts of quarterly changes—as distinct from the levels—of the variables.

(2) In both the predictions and the forecasts, root mean square errors (RMSE's) in such aggregates as GNP and personal consumption expenditures are substantially less than the sum of RMSE's in the components, indicating offsetting of errors.

(3) The errors in the post-sample predictions are generally much larger than in the sample period predictions. For trend-type variables, the excess of post-sample over sample period errors is less in percentage than in absolute terms, but still substantial.

(4) The *ex-ante* forecast errors are smaller than the post-sample prediction errors for real output, but substantially larger for the price level. It is concluded tentatively that judgmental adjustments in the constant terms helped to improve the forecasts.

(5) There is virtually no bias in the sample period predictions. However, errors in both the post-sample predictions and the forecasts show evidence of significant bias, though the significance tests are tenuous. The post-sample predictions show significant positive bias in real GNP and in certain GNP components, but not in the price level for the first few quarters. In the *ex-ante* forecasts, by contrast, bias in real GNP is positive, though not statistically significant until the fifth quarter. There is a strong negative bias in the price level forecasts, which offsets the positive bias in real output and results in relatively small errors in money GNP.

(6) For current-dollar GNP, *ex-post* forecasts (those in which the projected values of exogenous variables used in the *ex-ante* forecasts are replaced by actual values) yield larger errors than *ex-ante* forecasts for the whole forecast horizon; this results from offsetting errors in exogenous variables and model

(Continued on page 52)

(Continued from page 8)

months ending in June. An official of the United Auto Workers has stated that the union's demands in bargaining on the auto industry contract, which expires September 15, could shift to placing a much higher priority on getting a large wage increase, if the members begin to press for it. Recent statements by both the Treasury Secretary and the President of the AFL–CIO have acknowledged that if the price rise does not slow, labor can be expected to demand substantially larger wage increases later this year and in 1974 than has been the pattern thus far in 1973.

National Accounts in the Second Quarter

On the basis of more complete source data, BEA has revised the estimate of second quarter GNP. The revisions of the components are very small, and have a negligible effect on total GNP. Real GNP is very slightly smaller than previously estimated, current dollar GNP is slightly larger, and the implicit price deflator is also higher.

Corporate profits

According to the preliminary estimate, corporate book profits before taxes increased $10½ billion in the second quarter to a seasonally adjusted annual rate of $130.1 billion. Book profits had risen $13¾ billion in the first quarter, and about $19 billion in the year from end-1971 to end-1972. More than one-third of the second quarter increase was in durable goods manufacturing; advances were also reported by nondurables manufactures and financial institutions.

Book profits include gains or losses due to differences between the replacement cost of goods taken out of inventory and the cost at which these items are charged to production. Profits as measured in the national income accounts (NIA) include only profits arising from current production, and thus exclude inventory profits—which were at an annual rate of $21 billion in the second quarter, up $5¾ billion from the first. Thus, pretax profits on the national income basis rose $4¾ billion to a seasonally adjusted annual rate of $109 billion. The national income profits figure had risen $5½ billion in the first quarter and $15½ billion from end-1971 to end-1972.

Federal fiscal position

The Federal fiscal position as measured in the NIA was in approximate balance in the second quarter. Large increases in receipts far outpaced the growth of expenditures, and the accounts registered a nominal surplus of $.1 billion at a seasonally adjusted annual rate. There has been deficits of $5 billion in the first quarter and $23.4 billion in the fourth; a surplus was last registered in the fourth quarter of 1969.

For the fiscal year 1973, the preliminary figures (based on data not seasonally adjusted) show a Federal deficit of $12.1 billion as measured in the NIA. This compares with a $26.6 billion deficit estimated on the basis of the January budget, with NIA receipts having come in higher, and expenditures lower, than indicated in the budget. Table 3 shows the relationship between the unified budget and the NIA measures of receipts and expenditures, as estimated in the January budget and as indicated by the preliminary actual figures. Detailed explanations of the reconciliation items appear on page 28 of the February 1973 issue of the SURVEY.

Table 3.—Relationship of Federal Government Receipts and Expenditures in the National Income Accounts to the Unified Budget, Fiscal Year 1973

[Billions of dollars]

	January 1973 budget	August 1973 preliminary
Receipts		
Unified budget receipts	225.0	232.2
Coverage differences	–.2	–.3
Netting and grossing	5.3	6.0
Timing differences	2.6	5.2
Miscellaneous	–.4	–.1
Federal receipts, NIA basis	232.3	242.9
Expenditures		
Unified budget outlays	249.8	246.4
Coverage differences	.6	.3
Outer continental shelf deposit funds	1.1	1.1
Other	–.5	–.8
Financial transactions	–.9	–1.8
Net purchases of land	2.9	2.8
Netting and grossing	5.3	6.0
Timing differences	2.5	1.6
Miscellaneous	–.6	.3
Federal expenditures, NIA basis	258.9	255.0
Unified budget surplus or deficit (–)	–24.8	–14.2
NIA deficit surplus or deficit (–)	–26.6	–12.1

Source: Estimates by BEA.

(Continued from page 38)

equations. Differences between errors in *ex-post* and *ex-ante* forecasts of real GNP, the private GNP implicit deflator, and the unemployment rate are generally slight.

(7) For key aggregates—current-dollar GNP, real GNP, real personal consumption expenditures, and the unemployment rate—model forecasts are superior to extrapolations based on autoregressions; for the private GNP deflator, some GNP components, and personal income, the latter are superior.

(8) While the model tends to underestimate the amplitude of cyclical swings in real output, it generally recognizes turning points. However, accuracy in identifying the quarter of downturn or upturn diminishes as the prediction horizon lengthens.

U. S. GOVERNMENT PRINTING OFFICE : 1973 O - 511-688